National Park Service
U.S. Department of the Interior

Natural Resource Program Center

Integrity of benthic macroinvertebrate communities in Bluestone National Scenic River, Gauley River National Recreation Area, and New River Gorge National River

Eastern Rivers and Mountains Network 2009 summary report

Natural Resource Data Series NPS/ERMN/NRDS—2010/026

ON THE COVER
Bucklick Branch at New River Gorge National River.
Photograph by: Caleb Tzilkowski.

Integrity of benthic macroinvertebrate communities in Bluestone National Scenic River, Gauley River National Recreation Area, and New River Gorge National River

Eastern Rivers and Mountains Network 2009 summary report

Natural Resource Data Series NPS/ERMN/NRDS—2010/026

Caleb J. Tzilkowski, Kristina K. Callahan, Matthew R. Marshall, and Andrew S. Weber

National Park Service
Northeast Region
Eastern Rivers and Mountains Network
Forest Resources Building
University Park, PA 16802

February 2010

U.S. Department of the Interior
National Park Service
Natural Resource Program Center
Fort Collins, Colorado

The National Park Service, Natural Resource Program Center publishes a range of reports that address natural resource topics of interest and applicability to a broad audience in the National Park Service and others in natural resource management, including scientists, conservation and environmental constituencies, and the public.

The Natural Resource Data Series is intended for timely release of basic data sets and data summaries. Care has been taken to assure accuracy of raw data values, but a thorough analysis and interpretation of the data has not been completed. Consequently, the initial analyses of data in this report are provisional and subject to change.

All manuscripts in the series receive the appropriate level of peer review to ensure that the information is scientifically credible, technically accurate, appropriately written for the intended audience, and designed and published in a professional manner. Data in this report were collected and analyzed using methods based on established, peer-reviewed protocols and were analyzed and interpreted within the guidelines of the protocols.

This report is available from the Eastern Rivers and Mountains Network website (http://science.nature.nps.gov/im/units/ERMN) and the Natural Resource Publications Management website (http://www.nature.nps.gov/publications/NRPM).

Please cite this publication as:

Tzilkowski, C. J., K. K. Callahan, M. R. Marshall, and A. S. Weber. 2010. Integrity of benthic macroinvertebrate communities in Bluestone National Scenic River, Gauley River National Recreation Area, and New River Gorge National River: Eastern Rivers and Mountains Network 2009 summary report. Natural Resource Data Series NPS/ERMN/NRDS—2010/026. National Park Service, Fort Collins, Colorado.

NPS 619/101086, 600/101086, 637/101086, February 2010

Contents

Figures

Page

Tables

Page

Abstract

During 2008, the Eastern Rivers and Mountains Network (ERMN) of the National Park Service (NPS) began monitoring benthic macroinvertebrate (BMI) communities in wadeable streams throughout its nine parks. There were 26 randomly selected sites and seven “targeted” (i.e., non-random) sites sampled throughout Bluestone National Scenic River (BLUE), Gauley River National Recreation Area (GARI), and New River Gorge National River (NERI) during March and early April 2009. In addition to BMI samples, core water quality data (i.e., temperature, dissolved oxygen, pH, and specific conductance) were collected and reach-scale habitat was characterized.

Core water quality parameters at BLUE, GARI, and NERI sites were typical of forested watersheds with similar geologic characteristics. Relationships among core parameters were also typical – specific conductance generally decreased with decreasing pH. The pH of BLUE, GARI, and NERI streams ranged from 5.59 (Richlick Branch) to 8.58 (Arbuckle Creek), but the vast majority of sites had pH between 7.00 and 8.00. In addition to having high pH (8.53), water at the Bluestone River sites had, by far, the greatest specific conductance (≈ 425 µS/cm) of any of the sampled sites.

Benthic macroinvertebrate communities throughout BLUE streams had Macroinvertebrate Biotic Integrity Index (MBII) values that ranged from 34.1 (Mountain Creek) to 48.5 (Bluestone River [Tramway]). Based on MBII thresholds for the Southern Appalachians Ecoregion, none of the sampled BLUE sites were considered to be in the “Good” condition class. One site (Mountain Creek) was estimated to be in “Poor” condition whereas the remaining three sites were in the “Fair” condition class. At GARI, all stream condition classes were represented among sampled streams; each class was represented by one stream. Based on the MBII, Horseshoe Creek ranked highest (MBII = 54.2), followed by Meadow Creek (46.5), and then by Laurel Creek (34.0), which was considered to be in Poor condition. Benthic macroinvertebrate communities throughout NERI had MBII values that ranged from 16.3 (Wolf Creek) to 73.7 (Laurel Creek). There were eight NERI stream sites considered to be in Good condition, 11 sites were in Fair condition, and 6 sites were in Poor condition.

Based on the probabilistically chosen sites, distribution of wadeable GARI and NERI streams among condition classes was better than streams throughout the ecoregion as reported in the U.S. Environmental Protection Agency’s Wadeable Streams Assessment (WSA). As detailed in the report, methodological differences between the ERMN and WSA likely led to underestimates of BLUE, GARI, and NERI stream condition. What this means is, in the framework of the WSA, ecological integrity of GARI and NERI wadeable streams was likely even greater than we estimated during 2009. With each future sampling season, the ERMN BMI monitoring program will be refined and improved, which will allow more precise and accurate comparisons to be made among BLUE, GARI, and NERI streams and streams throughout the region.

Introduction

During 2008, the Eastern Rivers and Mountains Network (ERMN) of the National Park Service (NPS) began monitoring benthic macroinvertebrate (BMI) communities in wadeable streams throughout its nine parks. This monitoring effort is a component of the ERMN Vital Signs monitoring program (Marshall and Piekielek 2007) as part of the nationwide NPS Inventory and Monitoring Program (Fancy et al. 2009).

One of the primary objectives of the ecological monitoring program in the ERMN is to evaluate status and trends in the condition of tributary watersheds flowing into and through member parks. Watershed condition is evaluated using measures of ecosystem integrity, including streamside bird species and communities (Mattsson and Marshall 2009), forest structure and composition (Perles et al. 2009), stream-dwelling benthic macroinvertebrates (Tzilkowski et al. 2009), stream chemistry, and watershed land use, type, and configuration (Marshall and Piekielek 2007). A primary purpose of the benthic macroinvertebrate monitoring protocol is to support the antidegradation or restoration of ERMN aquatic communities and their habitat (including water quality) by communicating monitoring program results to appropriate regulatory state and federal agencies.

Benthic macroinvertebrates are aquatic invertebrate animals larger than microscopic size that live on or within the stream bottom (benthos), and because they are a vital component of all functioning stream ecosystems, they are often used as indicators of ecosystem integrity. Types of BMI that are commonly used for water quality assessment include arthropods (insects, arachnids, and crustaceans), worms, clams, and snails. In addition to being instrumental to nutrient and carbon dynamics, BMI are an important link between basal resources (e.g., algae and detritus) and higher trophic levels (e.g., fish and birds) in stream food webs. Because BMI have been by far the most commonly used group for biological monitoring of aquatic ecosystems (Carter and Resh 2001), many metrics have been evaluated with respect to natural variation and responses to various sources of human-induced degradation. Given the proven ability to derive ecosystem integrity based on measures of BMI assemblage structure and composition, combined with the relatively low cost to sample, BMI are almost certainly the single best biological group to assess and monitor the ecological integrity of small and mid-sized streams.

At the time that this report was prepared, the BMI-monitoring protocol (Tzilkowski et al. 2009) had been developed, written, and received internal peer review but had not undergone the final peer review process. This report was intended to provide preliminary results to natural resource managers at Bluestone National Scenic River (BLUE), Gauley River National Recreation Area (GARI), and New River Gorge National River (NERI), and at cooperating entities. The preliminary nature of data presented in this report should be considered prior to its use or dissemination.

Methods

Although a brief overview of the BMI monitoring methods is provided here, a detailed rationale of the sampling design and methods, in addition to Standard Operating Procedures, are provided in the BMI Monitoring Protocol (Tzilkowski et al. 2009). Much of this protocol is based on protocols developed by the U.S. Geological Survey ([USGS] Moulton et al. 2000, Moulton et al. 2002) and Bowles et al. (2006) because those protocols and programs have already undergone considerable evaluation and revision. We modified those protocols to fit the character of ERMN parks and anticipated monitoring resources.

Site Selection

There are two types of sampling sites in the BMI Monitoring Program – probabilistic (i.e., stratified-random) sites and non-random "targeted" sites. The probability-based design was developed by Mattsson and Marshall (2009) for the ERMN Streamside Bird Monitoring Program. This design was adopted for the BMI Monitoring Program for several reasons: (1) the design provided a population of wadeable stream sizes (i.e., generally 2nd to 4th Strahler stream order) that were suited to sampling methods and metrics that have been thoroughly developed and tested; (2) the population of "medium-sized" streams were more hydrologically stable than smaller intermittent streams and more safely and consistently accessible than larger rivers; and (3) collocation of BMI sites with Streamside Bird and water quality monitoring sites will provide multiple lines of evidence, with both terrestrial and aquatic components, to better evaluate trends in ecosystem condition at a landscape scale. Targeted site locations were chosen for parks (e.g., BLUE) that did not suit the random design or in situations that were of particular interest (e.g., Glade Creek). Targeted sites were chosen in consultation with Jesse Purvis and Lisa Wilson (natural resource management staff at the parks).

There were 26 randomly selected sites and seven "targeted" sites throughout the three parks (Table 1). The four targeted sites at BLUE were collocated with water quality sites maintained by park staff – two of those sites (Little Bluestone River and Mountain Creek) were also Streamside Bird monitoring locations (Figure 1). Peters Creek was the only targeted site at GARI and was chosen to supplement the parks' water quality monitoring in that stream – the other three GARI sites were randomly selected and collocated with Streamside Bird monitoring sites (Figure 2). The vast majority of sites (n = 23) at NERI (Figure 3) were randomly selected and collocated with Streamside Bird monitoring sites, but two relatively large streams (Glade Creek and Piney Creek) had targeted sites located on them because of their size and resultant ecological and societal importance at NERI.

Field Methods

The sampling unit for the BMI monitoring program is the stream reach, which, for the ERMN program is defined as a length of stream chosen to represent a uniform set of physical, chemical, and biological conditions within a stream segment. The length of sampled reaches differs among watersheds but their length is proportional (i.e., 40 x) to stream width. Minimum and maximum reach lengths are 150 m and 500 m, respectively. Tributary reaches within floodplains of large rivers (e.g. New River) were typically not considered for sampling because those sites were thought to exhibit considerable natural variation due to the river itself.

Table 1. Types of benthic macroinvertebrate monitoring sampling sites throughout Bluestone National Scenic River (BLUE), Gauley River National Recreation Area (GARI), and New River Gorge National River (NERI). Probabilistic sites were chosen with a stratified-random design and collocated with Streamside Bird monitoring sites whereas "Targeted" sites were non-randomly located.

Park	Stream	Probabilistic sites	Targeted sites
BLUE	Bluestone River (Mouth)		X
	Bluestone River (Tramway)		X
	Little Bluestone River		X
	Mountain Creek		X
GARI	Horseshoe Creek	X	
	Laurel Creek (GARI)	X	
	Meadow Creek (GARI)	X	
	Peters Creek		X
NERI	Arbuckle Creek	X	
	Batoff Creek	X	
	Big Branch	X	
	Bucklick Branch	X	
	Buffalo Creek	X	
	Camp Branch	X	
	Davis Branch	X	
	Dowdy Creek	X	
	Ephraim Creek	X	
	Fall Branch	X	
	Fire Creek	X	
	Glade Creek		X
	Keeny Creek	X	
	Laurel Creek (NERI)	X	
	Little Laurel Creek	X	
	Meadow Creek (NERI)	X	
	Mill Creek	X	
	Piney Creek		X
	Richlick Branch	X	
	Slater Creek	X	
	UNT to Buffalo Creek	X	
	UNT to Laurel Creek (Backus Mtn)	X	
	UNT to Laurel Creek (Highland Mtn)	X	
	UNT to Meadow Creek	X	
	Wolf Creek	X	

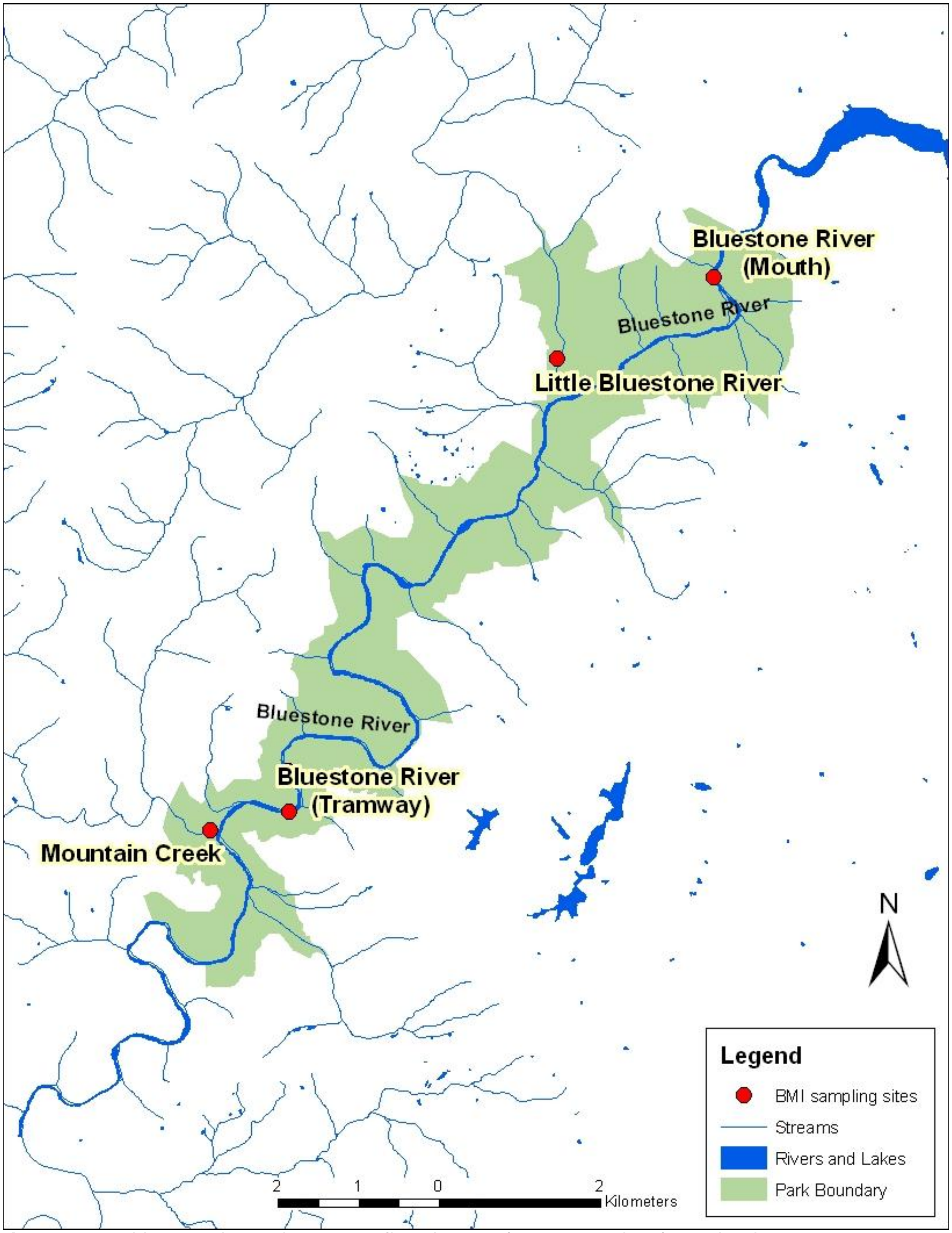

Figure 1. Benthic macroinvertebrate sampling sites at Bluestone National Scenic River.

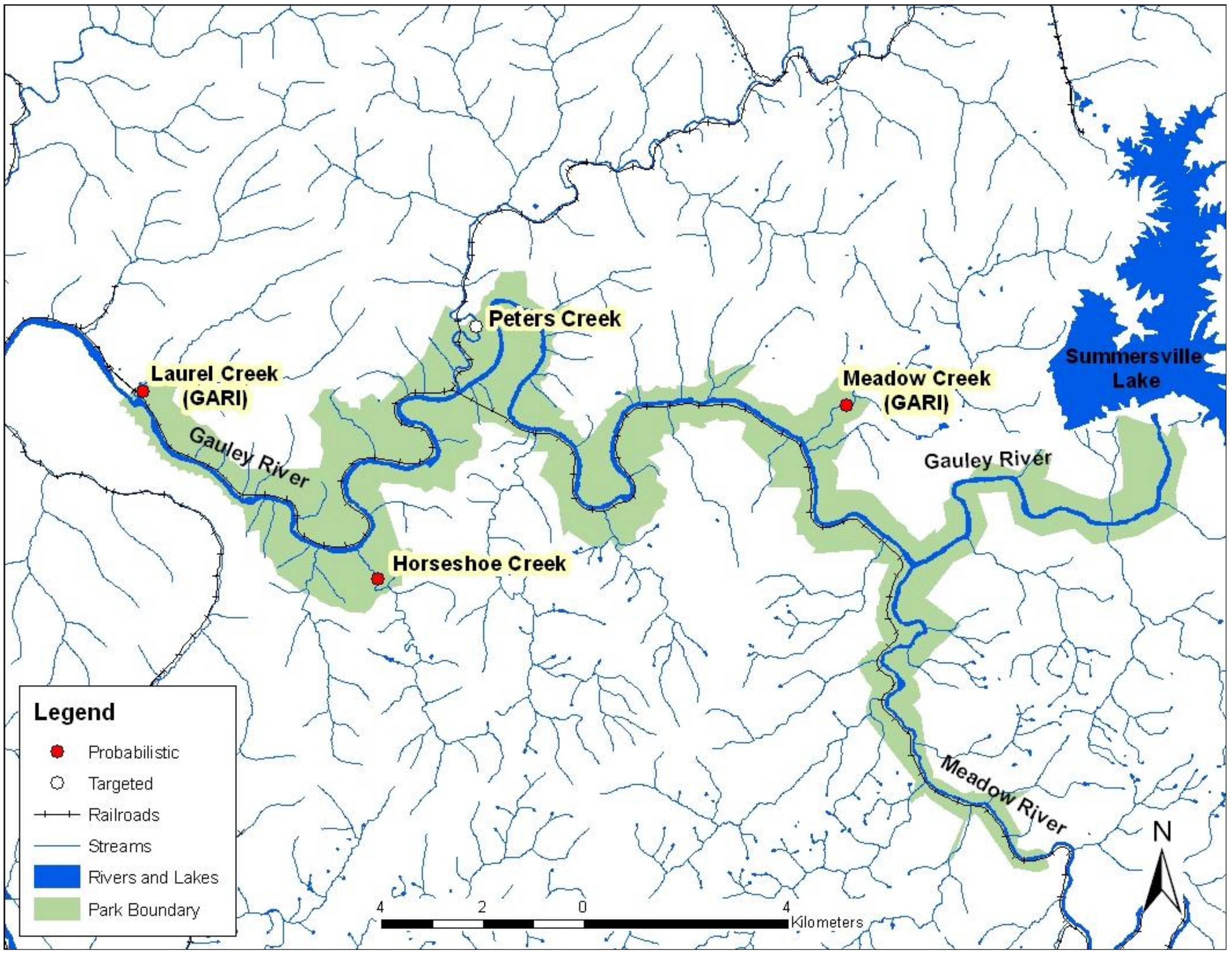

Figure 2. Benthic macroinvertebrate sampling sites at Gauley River National Recreation Area.

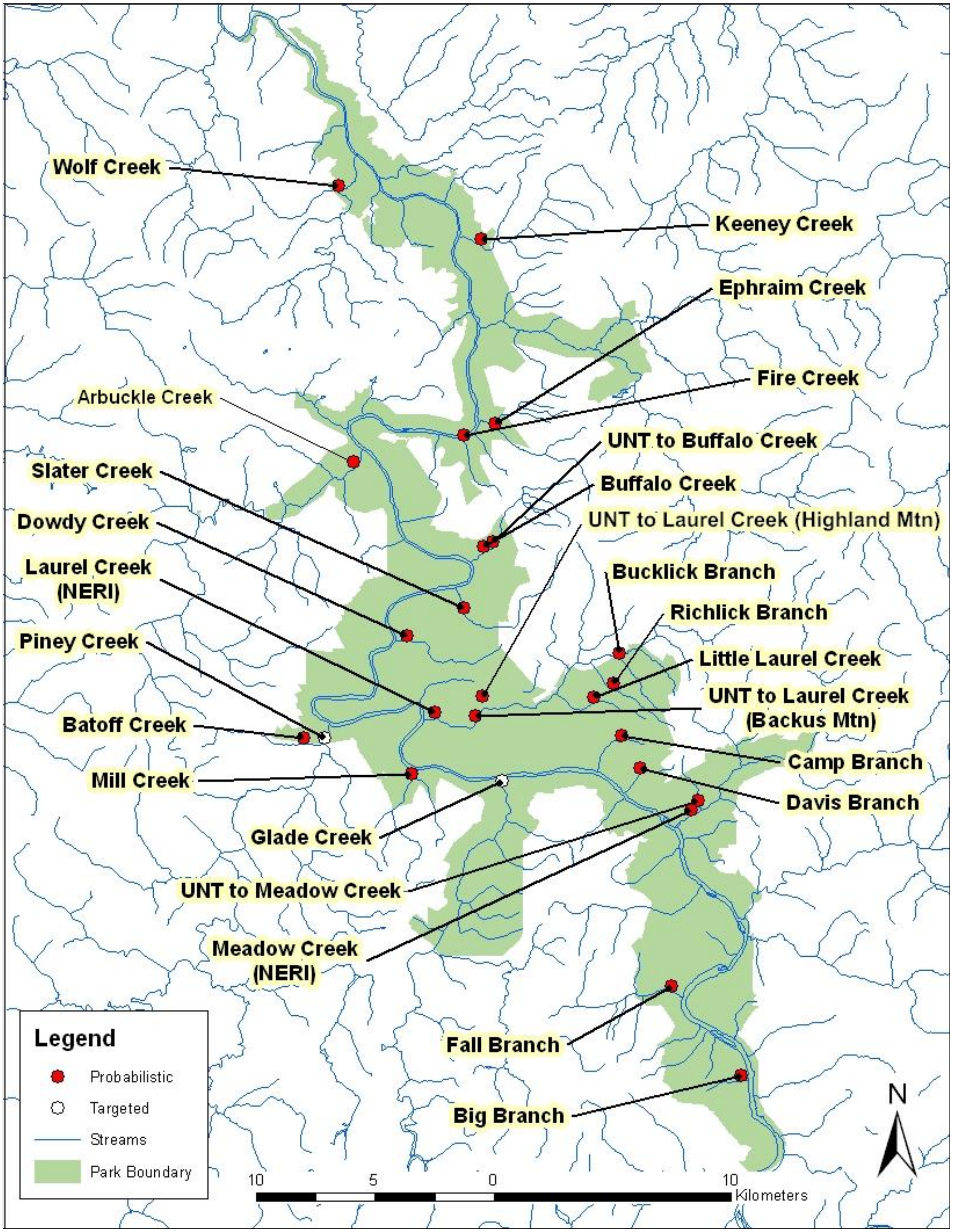

Figure 3. Benthic macroinvertebrate sampling sites at New River Gorge National River.

Because the probabilistic site selection strategy was developed for the Streamside Bird Monitoring Program, these BMI sites were collocated with Streamside Bird monitoring sites and anchored on the point count location that is closest to the point of access (either upstream-most or downstream-most point count). Sampling is conducted within Streamside Bird transects; therefore, sampling progresses either upstream or downstream into transects, depending on whether the anchor point is at the upstream- or downstream-most end of the transect.

Although sampling was proposed to be conducted during the fall, it was apparent from sampling efforts during 2007 (water sampling; Caleb Tzilkowski, personal observation) and 2008 (pilot BMI sampling; Caleb Tzilkowski, personal observation) that many "medium-sized" streams throughout the parks are "dry" too frequently during the fall for a long-term monitoring program. Virtually all of the randomly selected sites lacked surface flow during October 2008, and although those exact locations were not visited in 2007, similarly sized streams at the parks were dry then as well. Consequently, it was decided that sampling at BLUE, GARI, and NERI will be conducted during the spring (i.e., March–April) into the future. During 2009, sampling was conducted from March 23rd through April 1st, which worked well. Only one site (Peters Creek; GARI) slated to be sampled was not sampled during 2009 due to a combination of logistical and weather issues. We were unable to return to GARI to sample Peters Creek before a large storm in early April elevated stream levels to the extent that we would have had to return weeks later. The two Bluestone River sites were sampled during October 2008, but high flows prevented them from being sampled during spring 2009.

The ERMN method typically used for collecting BMI is termed semi-quantitative richest-targeted habitat (RTH) sampling, which is a type of disturbance-removal sampling (Moulton et al. 2002). Although similar to more common kick sampling methods, RTH sampling calls for consistent and thorough collection of BMI from a fixed area; thus, it is considered a more precise method and allows for estimation of stream productivity unlike many other sampling methods. Many BMI disturbance-sampling methods are qualitative (not quantitative) and are comparatively inconsistent because there is no measurement of sampling area – instead, those methods usually rely on a timed sampling effort. For the RTH method, five discrete samples are collected from riffles throughout the reach and are ultimately composited into a single homogenous sample. Ideally, discrete samples are taken from different riffles, but if fewer than five riffles are present, samples may be taken from the same riffle. Physical conditions (i.e., depth, flow, and substrate) are recorded at each sampling location and should be as similar as possible among replicates. Sampling is conducted by defining a 0.25 m^2 sampling area with a template and then disturbing substrate within that area so that BMI are dislodged and then drift into a net placed downstream of the sampling area. The composited samples result in 1.25 m^2 of sampled area at each site.

The RTH method is not designed for sampling large streams because near mid-channel (where sampling is preferred), large streams are typically too deep and have large (e.g., boulder) substrate. Several of the larger NERI streams (Arbuckle Creek, Glade Creek, Piney Creek, and Wolf Creek) and the two mainstem Bluestone River sites were difficult to sample with the RTH. Using the RTH method at those sites will likely cause sampling effectiveness (thus assessments) to be variable among years. To overcome those difficulties, we are considering deploying multiplates (artificial substrate sampling devices) in the future for several weeks during March and April at these sites. Multiplates provide substrate for BMI colonization for a fixed exposure period, after which the sampler is retrieved and the attached organisms are harvested. Using multiplates allows comparison of results from different locations and times by providing uniformity of substrate type, depth, and exposure duration (Bode et al. 2002). It should be noted that BMI that colonize multiplates are influenced more by water quality than by stream bottom conditions; consequently, comparisons among multiplate samples and other sampling gear (e.g., kick nets) should be done with caution. For these reasons, we will continue to sample using the RTH method in addition to deploying multiplates at these sites (if logistics allow) for several years for comparative purposes prior to making a final decision on sampling method at these sites.

In addition to BMI samples, core water quality data (i.e., temperature, dissolved oxygen [DO], pH, and conductivity) were collected and reach-scale habitat was characterized using the U.S. Environmental Protection Agency rapid bioassessment method (Barbour et al. 1999). Samples were processed in the field by using an elutriation method to remove mineral materials and large organic matter (e.g., whole leaves and sticks). Samples were preserved in 95% ethanol, packed carefully, and transported back to the laboratory for processing and identification.

Laboratory Methods

Laboratory methods for processing samples in the ERMN BMI Program rely a great deal on procedures developed by the USGS (Moulton et al. 2000). A fixed-count subsample of 300 ± 20% individuals are sorted and identified from each sample. The relatively large subsample size yields data that meets quality standards (i.e. precision and accuracy) required by most monitoring programs; however, processing and identifying additional individuals (> 300) does not typically yield enough additional information to justify the added effort (Moulton et al. 2000). Generally, BMI were identified to genus using standard dichotomous keys, but some groups (e.g., Chironomidae, Oligochaeta) were identified to coarser taxonomic levels. Microsoft Access 2007 is the primary software used for storing and managing ERMN BMI and stream habitat data, whereas the Invertebrate Data Analysis System (IDAS *version 5*, U.S. Geological Survey, Raleigh, NC) was used for resolving taxonomic ambiguity issues and calculating metrics that describe the structure and diversity of BMI communities.

Data Analysis

We calculated all BMI community metrics with IDAS and calculated the Macroinvertebrate Biotic Integrity Index (MBII; Klemm et al. 2003) using Microsoft Excel 2007. The MBII was developed by the U.S. Environmental Protection Agency's (USEPA) Environmental Monitoring and Assessment Program (EMAP) and was ultimately used for the USEPA's Wadeable Stream Assessment (WSA, USEPA 2006, Herlihy et al. 2008).

The rationale behind biotic integrity indices is that a suite of metrics that represent community structure, pollution tolerance, functional feeding groups and habitat occurrences, life history strategies, disease, and density provide insights regarding how biological communities respond to different natural and anthropogenic stressors (Klemm et al. 2003). A common stream bioassessment practice is to compare BMI community metrics from candidate streams to the same metrics from reference streams. Reference streams are "least disturbed," similarly sized streams within comparable geographic and geologic settings that provide an estimate of least-impaired stream communities. Departure of the sampled BMI community from expected BMI community composition (i.e., reference streams) serves as a measure of stream impairment. The MBII is one such index that uses reference streams to assess stream impairment.

The MBII was chosen for use in the ERMN because it was developed for upland and lowland streams dominated by riffle habitat in the Mid-Atlantic Highlands Region (MAHR). Moreover, the MBII was based on a large dataset of 574 wadeable stream reaches and was thoroughly tested. The MBII is a broadly applicable measure of stream impairment because it is based on several factors that affect aquatic communities throughout the MAHR. Impaired and reference streams for the MBII were identified by Klemm et al. (2003) using water chemistry, qualitative habitat, and minimum organism count criteria. Impaired reaches were defined by meeting any one of the following criteria: pH <5, chloride >1000 µeq/L, sulfate >1000 µg/L, total phosphorous >100 µg/L, total nitrogen >5000 µg/L, or a mean qualitative habitat score <10 (of a possible 20). Reference reaches met all of the following criteria (Klemm et al. 2003): sulfate <400 µg/L, Acid Neutralizing Capacity (ANC) >50 µeq/L, chloride <100 µeq/L, total phosphorous <20 µg/L, total nitrogen <750 µg/L, mean qualitative habitat score >15, and at least 150 organisms.

The MBII uses seven metrics selected from the 100 that are commonly used by governmental agencies throughout the MAHR. The metrics chosen were those that performed best in terms of range, precision, responsiveness to various human-induced disturbances, relationship to catchment area, and redundancy (Table 2; Klemm et al. 2003). Most MBII metrics are counts or proportions of taxa in the community that are characterized as tolerant or intolerant to human perturbations; however, one of the metrics (Macroinvertebrate Tolerance Index; MTI) is more complex because it incorporates values (0–10) for each taxon with respect to pollution tolerance, weighted by taxon abundance, and results in higher scores as the proportion of taxa tolerant to general pollution increases (Klemm et al. 2003). Pollution Tolerance Values (PTV) incorporated in the MTI were average tolerances to "various types of stressors" (Klemm et al. 2002).

Table 2. Macroinvertebrate Biotic Integrity Index (MBII) metric descriptions and their directions of response to increasing human perturbation (Response) from Klemm et al. (2003).

Metric	Description	Response
Ephemeroptera richness	Number of Ephemeroptera (mayfly) taxa	Decrease
Plecoptera richness	Number of Plecoptera (stonefly) taxa	Decrease
Trichoptera richness	Number of Trichoptera (caddisfly) taxa	Decrease
Collector-filterer richness	Number of taxa with a collecting or filtering-feeding strategy	Decrease
Percent non-insect individuals	Percent of individuals that are not insects	Increase
Macroinvertebrate Tolerance Index	$\sum_i p_i t_i$, where p_i is the proportion of individuals in taxon *i* and t_i is the pollution tolerance value (PTV) for general pollution	Increase
Percent five dominant taxa	Percentage of individuals in the five numerically dominant taxa	Increase

There are important qualifications that should be considered while interpreting the 2009 data with the MBII. We present MBII ranges from the WSA (Herlihy et al. 2008) as points of reference; however, it must be recognized that our sampling methods were similar but not identical to those used to develop (Klemm et al. 2003) or apply the MBII for the WSA (Herlihy et al. 2008). An often encountered difficulty among BMI monitoring or assessment programs is that comparisons are made among datasets that have been compiled by different researchers using different methods. This reality has become increasingly accepted recently (Carter and Resh 2001), but is unfortunately not always recognized.

We do not yet know the degree to which methodological differences influence the comparability of ERMN data to other studies that use the MBII (e.g., WSA). We speculate that MBII scores at BLUE, GARI, and NERI were lower than they would have been if we used identical methods employed for MBII development. To design the MBII, Klemm et al. (2003) indentified chironomid midges to genus, whereas chironomids were only identified to family for BLUE, GARI, NERI samples. What these differences likely lead to is lower richness scores for several taxa (especially Chironomidae), and because the MBII is largely influenced by richness metrics, the MBII scores may have been lower as well. Regardless of comparability to other studies, the MBII and its constituent metrics reflect the condition of BLUE, GARI, and NERI streams relative to each other and to themselves through time.

We also present three other commonly used BMI community metrics (taxa richness, Shannon's Diversity and Evenness) for comparison because they are likely to be familiar to most readers of this report. Taxa richness was the combined number of unique taxa (usually genera). Shannon's diversity and evenness were calculated with IDAS using formulae provided by Brower and Zar (1984), which were:

Shannon's Diversity (H`): information theory-based index that measures the "uncertainty" of a taxon selected at random from the community. High diversity is associated with high uncertainty and low diversity with low uncertainty. This index is the equivalent of the Brillouin's diversity index, but it is intended for use when the abundance data come from a random sample of the community or subcommunity.

$$H` = (N \log_{10} N - \Sigma\, n \log_{10} n)/N$$

Shannon's Evenness (J`): ratio of the observed Shannon diversity to the maximum possible diversity (that is, diversity when individuals are distributed as evenly as possible among the species). Like the Shannon diversity index, this measure is intended to be used when the abundance data come from a random sample or the community or subcommunity

$$J` = H`/H_{max}` \text{ where } H_{max}` = \log_{10} S$$

Abbreviations used in formulae: S = number of taxa in sample, n = abundance of an individual taxon, N = total number of individuals in sample.

Results

Benthic Macroinvertebrate Communities

Bluestone National Scenic River

As mentioned in the methods section, the Bluestone River sites could not be sampled during spring 2009 because high stream flows prevented access to suitable habitat. Results from fall 2008 are presented for those sites but not considered to be entirely comparable to the remaining sites that were sampled during spring. Benthic macroinvertebrate communities throughout BLUE streams had MBII values that ranged from 34.1 (Mountain Creek) to 48.5 (Bluestone River [Tramway]; Figure 4). Based on MBII thresholds for the Southern Appalachians Ecoregion (Herlihy et al. 2008), none of the sampled BLUE sites were considered to be in the "Good" condition class. One site (Mountain Creek) was estimated to be in "Poor" condition, whereas the remaining three sites were in the "Fair" condition class.

Although the Little Bluestone River site had the third lowest MBII score at BLUE, that sample and the highest ranking sample (Bluestone River [Tramway]) both contained 29 taxa, whereas Mountain Creek had the fewest (25; Table 3). Similarly, Little Bluestone River had more Ephemeroptera, Plecoptera, and Trichoptera taxa than any other sample but had a relatively low MBII value due to its low diversity, evenness, and poor MTI measures. At most sites, the proportional metrics (%Non-insects and %5 dominant) and Shannon diversity and evenness metrics generally responded as expected; with increasing MBII scores, the proportional and Shannon metrics decreased and increased, respectively. As expected, the MTI decreased with increasing MBII scores and ranged from 4.32 to 5.13. Densities of BMI at both Bluestone River sites were approximately three times as great (>3,400 m^{-2}) as the estimates for Little Bluestone River (1,190 m^{-2}) and Mountain Creek (967 m^{-2}; Figure 5).

Gauley River National Recreation Area

At GARI, all stream condition classes were represented among sampled streams; each class was represented by one stream. Based on the MBII, Horseshoe Creek ranked highest (MBII = 54.2), followed by Meadow Creek (46.5), and then by Laurel Creek (34.0), which was considered to be in Poor condition. Total taxa richness ranged from 26 (Laurel Creek) to 37 (Horseshoe Creek), and as evidenced by the fairly different MBII values, the identities and pollution tolerance of taxa and their relative abundance in the community varied among sites. Metrics that comprise the MBII and the Shannon measures largely responded in the expected direction with increasing MBII with one anomaly – Meadow Creek had considerably more collector-filterer taxa (13) than either of the other sites. Benthic macroinvertebrate densities were relatively low, but they were similar in GARI streams and ranged from 509 m^{-2} (Meadow Creek) to 604 m^{-2} (Laurel Creek).

New River Gorge National River

Benthic macroinvertebrate communities throughout NERI had MBII values that ranged from 16.3 (Wolf Creek) to 73.7 (Laurel Creek; Table 4). There were eight NERI stream sites considered to be in Good condition, 11 sites were in Fair condition, and 6 sites were in Poor condition. Interestingly, there was a rather wide margin between the highest MBII (73.7; Laurel Creek) and the next highest MBII (60.7; Davis Branch). A similar situation was apparent on the other end of the spectrum – the site with the lowest MBII (16.3; Wolf Creek) was estimated to be in considerably poorer condition than the next poorest site (29.3; Richlick Branch).

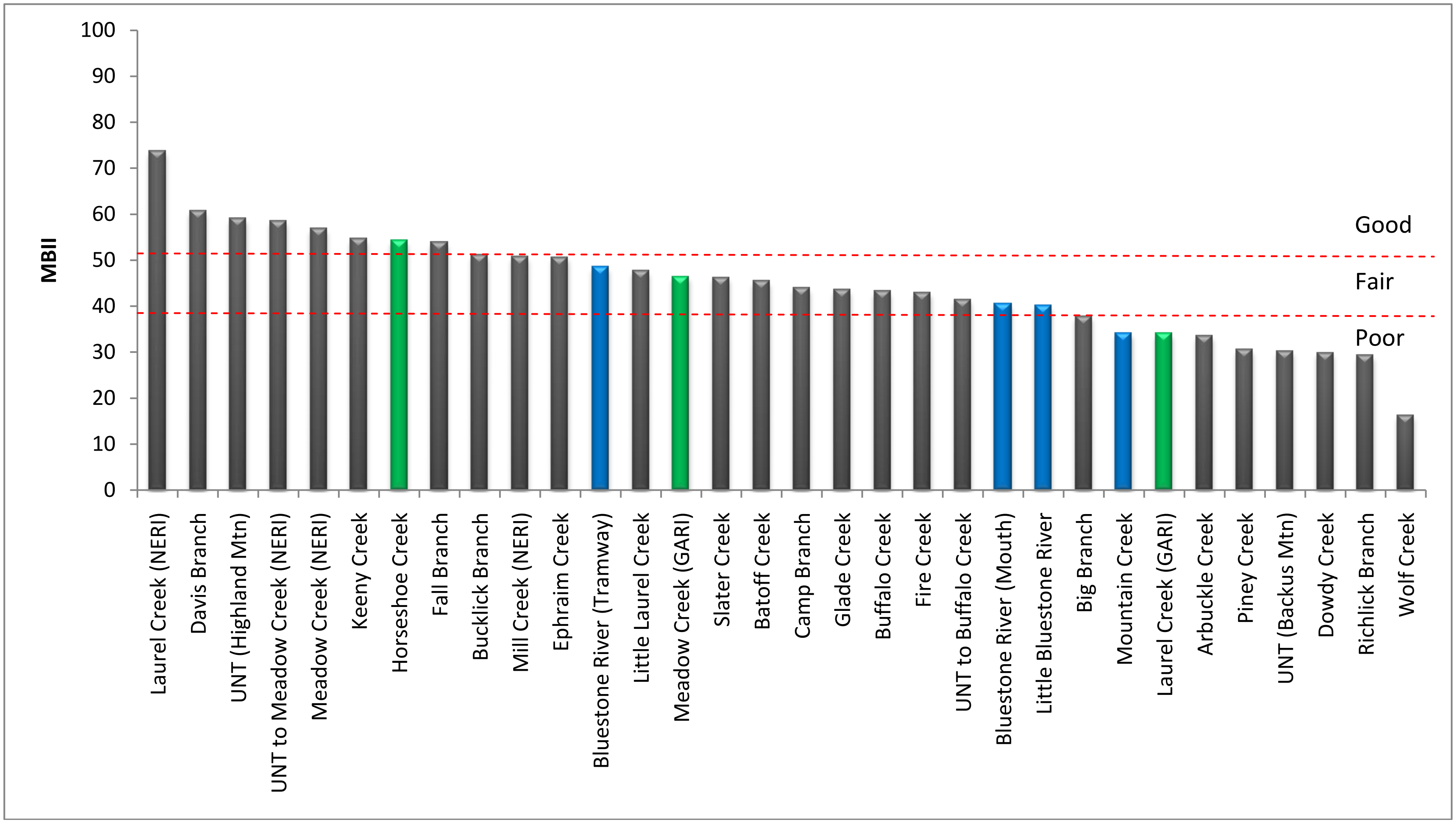

Figure 4. Macroinvertebrate Biotic Integrity Index (MBII, Klemm et al. 2003) values for benthic macroinvertebrate samples collected at sampling sites throughout Bluestone National Scenic River (blue bars), Gauley River National Recreation Area (green bars), and New River Gorge National River (gray bars). The two Bluestone River sites were sampled during October 2008 whereas all other sites were sampled in March 2009. Dashed red lines represent the 25th (51.0) and 5th (37.0) percentiles of index scores reported by Herlihy et al. (2008) for the Southern Appalachians Ecoregion as part of EPA's Wadeable Streams Assessment (WSA). These percentiles represented thresholds between "Good", "Fair", and "Poor" stream condition classes in the WSA.

Table 3. Summary metrics and multimetric indices for benthic macroinvertebrate communities sampled from Bluestone National Scenic River (BLUE), Gauley River National Recreation Area (GARI), and New River Gorge National River (NERI). The two Bluestone River sites were sampled during October 2008, whereas all other sites were sampled in March 2009. Direction of metric or index response to increasing stream ecosystem integrity are denoted parenthetically by + or -. Richness metrics included total taxa richness (Total), and richness of Ephemeroptera (E), Plecoptera (P), Trichoptera (T), and Collector or Filter feeders (C-F). Proportional metrics included the percent of individuals in samples that were non-insect taxa (%Non-insects) or that comprised the combined five dominant taxa in the community (%5 dominant). Indices were the Macroinvertebrate Tolerance Index (MTI) and the Macroinvertebrate Biotic Integrity Index (MBII).

	Richness (+)					**Proportional (-)**		**Shannon (+)**		**Indices**	
Stream	**Total**	**E**	**P**	**T**	**C-F**	**%Non-insects**	**%5 dominant**	**Diversity**	**Evenness**	**MTI (-)**	**MBII (+)**
Bluestone River (Tramway)	29	6	3	7	6	14	48	1.24	0.85	4.51	48.5
Bluestone River (Mouth)	26	5	4	7	5	7	65	1.08	0.76	4.76	40.5
Little Bluestone River	29	10	7	4	8	1	78	0.84	0.57	5.13	40.1
Mountain Creek	25	5	5	4	6	6	70	1.03	0.74	4.32	34.1
Horseshoe Creek	37	7	4	8	9	4	51	1.26	0.80	4.10	54.2
Meadow Creek (GARI)	32	7	3	6	13	5	63	1.17	0.78	4.58	46.5
Laurel Creek (GARI)	26	7	6	5	6	5	84	0.79	0.56	4.90	34.0

Table 4. Summary metrics and multimetric indices for benthic macroinvertebrate communities sampled from New River Gorge National River (NERI) during March 2009. Direction of metric or index response to increasing stream ecosystem integrity are denoted parenthetically by + or -. Richness metrics included total taxa richness (Total), and richness of Ephemeroptera (E), Plecoptera (P), Trichoptera (T), and Collector or Filter feeders (C-F). Proportional metrics included the percentage of individuals in samples that were non-insect taxa (%Non-insects) or that comprised the combined five dominant taxa in the community (%5 dominant). Indices were the Macroinvertebrate Tolerance Index (MTI) and the Macroinvertebrate Biotic Integrity Index (MBII).

	Richness (+)					**Proportional (-)**		**Shannon (+)**		**Indices**	
Stream	**Total**	**E**	**P**	**T**	**C-F**	**%Non-insects**	**%5 dominant**	**Diversity**	**Evenness**	**MTI (-)**	**MBII (+)**
Laurel Creek (NERI)	38	9	8	8	8	5	47	1.34	0.85	3.94	73.7
Davis Branch	37	7	7	9	9	1	60	1.21	0.77	3.79	60.7
UNT (Highland Mtn)	35	8	7	8	7	3	61	1.18	0.76	3.93	59.0
UNT to Meadow Creek (NERI)	33	8	6	5	9	2	53	1.23	0.81	3.44	58.5
Meadow Creek (NERI)	30	9	7	5	8	3	56	1.18	0.80	4.07	56.8
Keeny Creek	32	6	8	6	7	5	63	1.15	0.77	3.95	54.6
Fall Branch	30	8	6	4	9	2	59	1.19	0.80	3.63	53.8
Bucklick Branch	41	6	6	11	8	3	60	1.25	0.78	4.36	51.1
Mill Creek	31	8	5	5	9	1	63	1.15	0.77	3.91	50.7
Ephraim Creek	28	6	7	6	6	3	72	1.03	0.71	3.87	50.6
Little Laurel Creek	38	8	7	5	10	4	71	0.98	0.62	4.24	47.6
Slater Creek	23	5	7	5	4	2	81	0.85	0.62	3.59	46.2
Batoff Creek	24	4	6	7	5	0	84	0.86	0.62	4.34	45.4
Camp Branch	24	3	4	5	6	1	66	1.04	0.75	3.59	43.9
Glade Creek	28	8	5	6	8	9	66	1.13	0.78	4.15	43.5
Buffalo Creek	23	6	6	4	6	1	77	0.99	0.73	3.80	43.2
Fire Creek	22	4	6	6	3	2	82	0.86	0.64	3.89	42.9
UNT to Buffalo Creek	27	6	6	3	8	2	68	1.09	0.76	4.10	41.4
Big Branch	25	7	6	2	9	2	80	0.84	0.60	3.85	37.8
Arbuckle Creek	15	1	0	4	4	7	87	0.76	0.65	4.06	33.5
Piney Creek	21	2	3	7	3	6	93	0.50	0.38	5.55	30.5
UNT (Backus Mtn)	20	3	5	4	6	4	78	0.93	0.71	4.43	30.2
Dowdy Creek	22	1	4	6	6	11	75	0.98	0.73	4.27	29.8
Richlick Branch	22	4	4	3	7	2	84	0.87	0.65	4.07	29.3
Wolf Creek	11	1	1	2	4	1	100	0.08	0.08	5.30	16.3

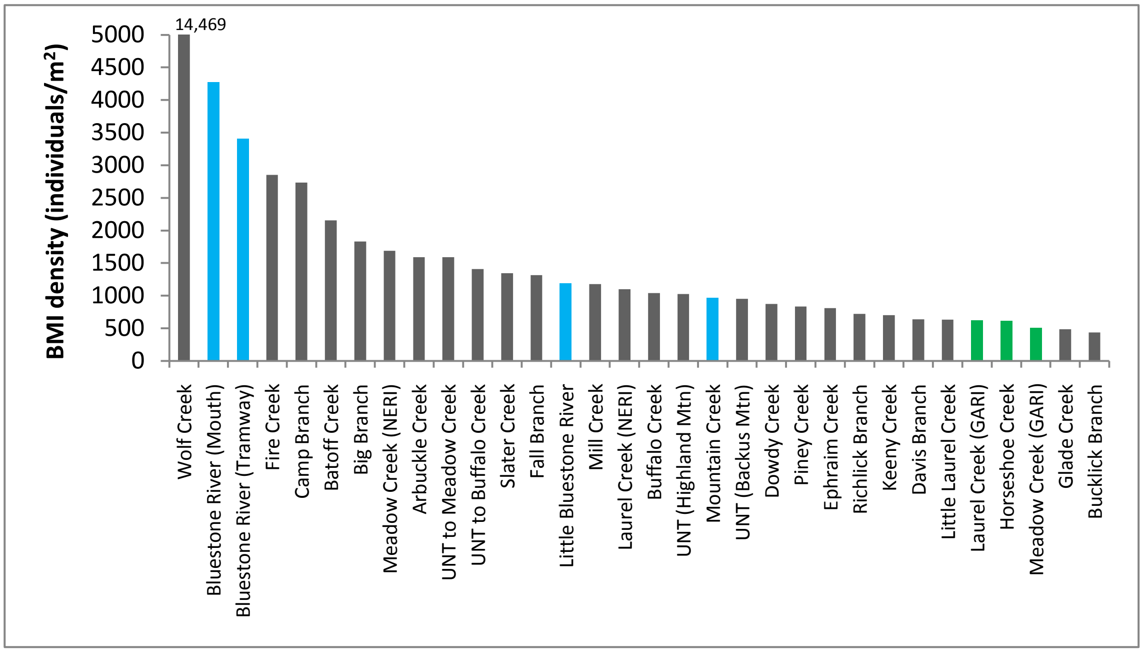

Figure 5. Density (individuals/m^2) of benthic macroinvertebrates collected at sampling sites throughout Bluestone National Scenic River (blue bars), Gauley River National Recreation Area (green bars), and New River Gorge National River (gray bars). The two Bluestone River sites were sampled during October 2008 whereas all other sites were sampled in March 2009. Note that the exceptionally great value for Wolf Creek (14,469 individuals/m^2) exceeds the axis maximum.

Total taxa richness in NERI samples ranged from four (Wolf Creek) to 38 (Laurel Creek). There were few surprises in the relationship between the MBII and individual metrics or the MTI. Interestingly, Bucklick Branch had considerably more Trichoptera taxa (11) than other streams despite being ranked 8th in terms of the MBII. Another interesting finding was that, throughout BLUE, GARI, and NERI, non-insects (e.g., worms, clams, snails) comprised a consistently low proportion of the BMI community. At NERI, the MTI generally increased with decreasing MBII, but there was one notable exception. Piney Creek had the poorest (i.e., highest) MTI score (5.55) despite being ranked 5th-poorest which would not be expected based solely on the pollution tolerance of taxa found there. Density of BMI was incredibly high in Wolf Creek (14,469 m-2) and was five times greater than the next highest density estimate (Fire Creek, 2,852 m-2). Otherwise, BMI densities ranged from 436 (Bucklick Branch) to 2,736 m-2 (Camp Branch).

Water Quality

Physical and chemical characteristics of water can vary markedly, both daily and annually. Although there are limitations to point-in-time characterizations of core water quality parameters, these measures can be helpful when evaluating patterns in biological data; moreover, extreme changes to these parameters can sometimes be detected with point-in-time samples. Generally, core water quality parameters (pH, specific conductance, temperature, DO) at BLUE, GARI, and NERI sites were typical of forested watersheds with similar geologic characteristics. Relationships among core parameters were also typical; specific conductance generally decreased with decreasing pH, whereas DO concentrations consistently decreased with increasing water temperature.

There were several notable observations regarding core water quality parameters at the three parks. The pH of BLUE, GARI, and NERI streams ranged from 5.59 (Richlick Branch) to 8.58 (Arbuckle Creek), but the vast majority of sites had pH between 7.00 and 8.00 (Figure 6). Dowdy Creek and Richlick Branch were the only streams with pH < 6.00, whereas Camp Branch and Bucklick Branch had pH between 6.00 and 7.00. The two Bluestone River sites and two NERI sites (Arbuckle Creek and Piney Creek) had pH greater than 8.3. In addition to having high pH, water at the Bluestone River sites had, by far, the greatest specific conductance (≈ 425 µS/cm) of any of the sampled sites (Figure 7). Given its odor, and personal communication with NERI staff, Arbuckle Creek was obviously polluted with sewage effluent, which was evidenced by the high specific conductance (280 µS/cm) observed there. Piney Creek and Wolf Creek suffer from similar effluent problems (Jesse Purvis, personal communication), as reflected by the relatively high specific conductance (≈ 180 µS/cm) at those sites. All other sites throughout the three parks had specific conductance ≤ 116 µS/cm; Richlick Branch, where there was obvious evidence of abandoned mine drainage, had the lowest specific conductance (20 µS/cm).

Water temperature throughout the parks was typical for the March sampling period and ranged between 5.0 (Mountain Creek) and 9.7^{o}C (Laurel Creek [NERI]; Figure 8). Although DO concentration was measured at all sites, the membrane on the DO meter began to operate improperly during the middle of the sampling effort and replacements were unavailable; consequently, the data were of limited value and not presented. We now have backup equipment so that this situation will not occur again. Given the spring sampling period (during the daytime) and consequent cold water, it is likely that even organically enriched streams would be nearly saturated, if not supersaturated, with oxygen during the daytime in March. Finally, as expected, there were no exceptionally warm or cold streams throughout BLUE, GARI, and NERI given the spring sampling period.

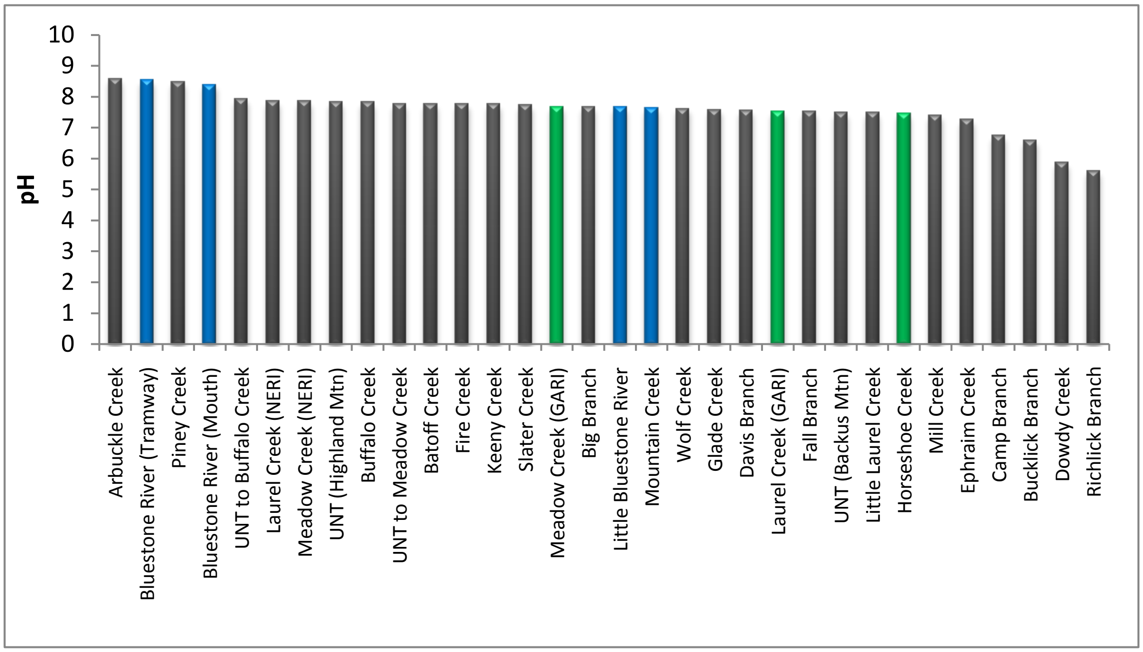

Figure 6. pH at sampling sites throughout Bluestone National Scenic River (blue bars), Gauley River National Recreation Area (green bars), and New River Gorge National River (gray bars). The two Bluestone River sites were sampled during October 2008 whereas all other sites were sampled in March 2009.

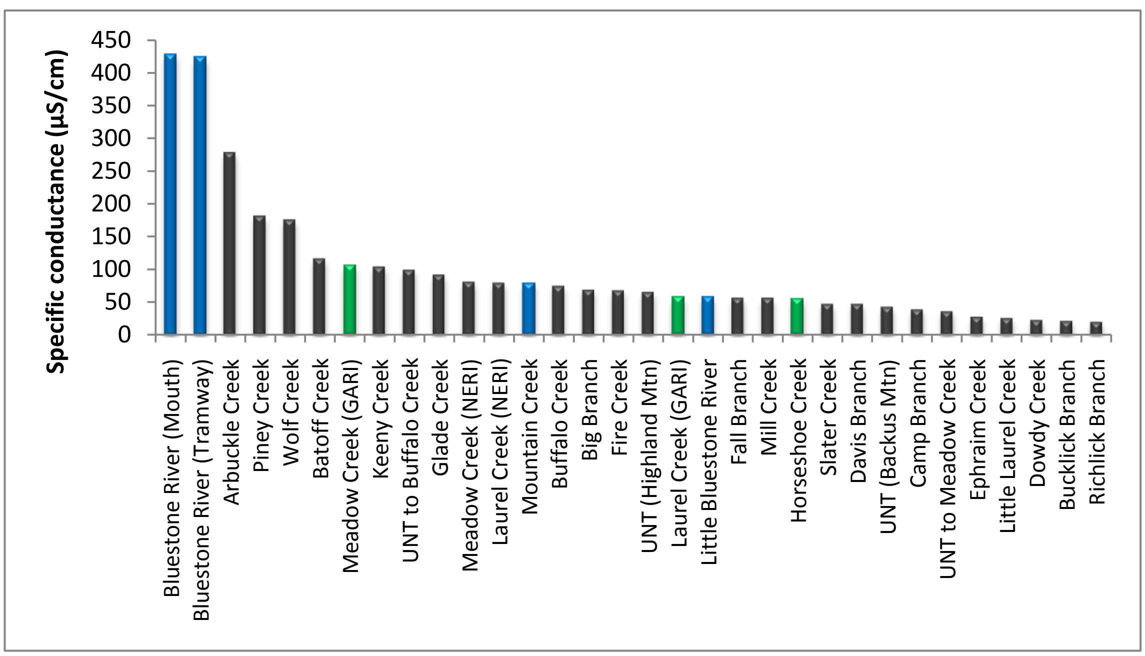

Figure 7. Specific conductance of water at sampling sites throughout Bluestone National Scenic River (blue bars), Gauley River National Recreation Area (green bars), and New River Gorge National River (gray bars). The two Bluestone River sites were sampled during October 2008, whereas all other sites were sampled in March 2009.

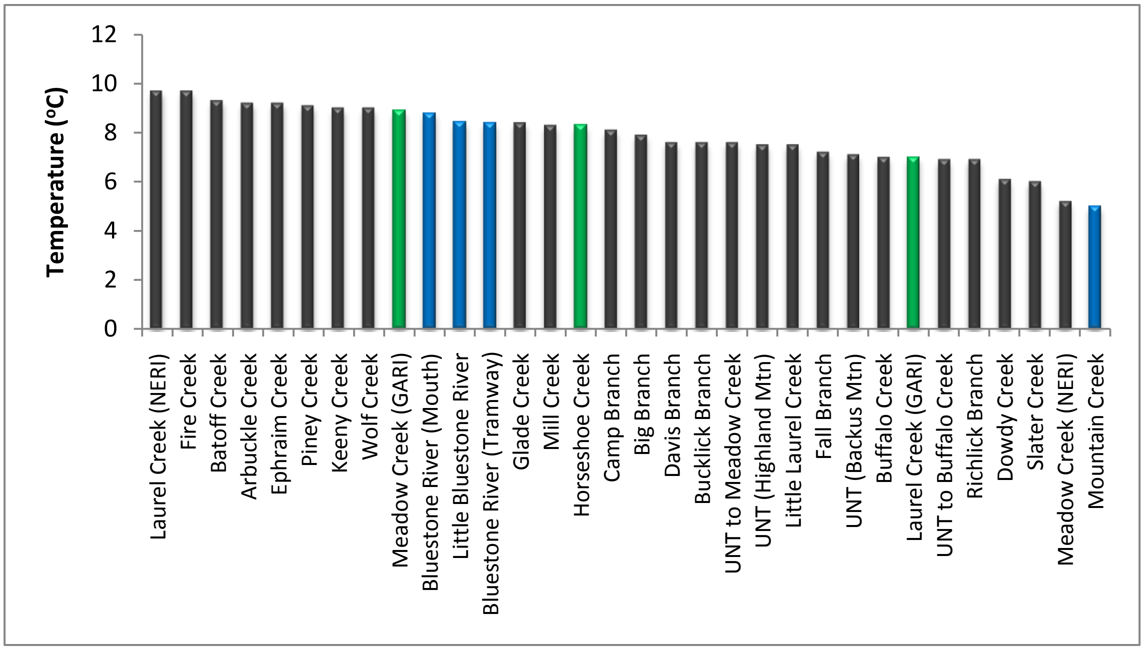

Figure 8. Temperature of water at sampling sites throughout Bluestone National Scenic River (blue bars), Gauley River National Recreation Area (green bars), and New River Gorge National River (gray bars). The two Bluestone River sites were sampled during October 2008 whereas all other sites were sampled in March 2009.

Discussion

This report summarized results from the first sampling season of the ERMN BMI monitoring program at BLUE, GARI, and NERI. The effort was largely successful in that it provided quality data for the majority of selected sites. All components of the protocol worked well, which was not a surprise because they were based largely on widely used USGS protocols. The primary challenge to interpreting the data (as discussed in the methods section) was that, because the ERMN protocol did not precisely follow other state or regional protocols, comparing our data with other efforts included qualifications.

It was originally proposed that all ERMN BMI sites would be sampled during the fall. It became apparent during protocol development that fall sampling across the network would not be possible for two reasons: 1) many high gradient streams in West Virginia (e.g., throughout GARI and NERI) often lack surface flow during the fall; and 2) logistical and personnel constraints would not allow all ERMN sites to be sampled during either of the ideal spring or fall sampling windows (i.e., October or March). Sampling conditions were ideal for much of the spring 2009 sampling season at BLUE, GARI, and NERI, although rain near the end of sampling elevated stream water levels and turbidity. Sampling will be coordinated in the future so that hydrologic conditions match those of 2009 to the extent possible. As discussed previously, we plan to deploy multiplate samplers at the relatively large, targeted sampling sites for at least the next several years because we are concerned that difficulty sampling deep, boulder habitat could lead to high sampling error. After comparing results from multiplate and RTH sampling at those sites, we will decide which method to permanently adopt in consultation with Jesse Purvis.

We compared ERMN results with results from the recently conducted USEPA Wadeable Streams Assessment (WSA). Decisions made during early stages of protocol development, which were discussed in the methods section, resulted in differences between the ERMN protocol and the WSA. A considerable difference was that the ERMN protocol calls for identification of chironomid midges to the family level, whereas the WSA used genus level identifications of that group. Our speculation was that the differences may have resulted in conservative estimates (i.e., underestimate) of WSA condition class (i.e., good, fair, poor) for BLUE, GARI, and NERI streams based on the MBII. Even if 2009 results were not considered an underestimate of stream condition class, condition of randomly chosen GARI and NERI streams was better than the broader Southern Appalachians Ecoregion (USEPA 2006; Figure 9). This statement is made based on only the probabilistically chosen sites, not the targeted (non-random) sites. Condition of targeted sites was either in the Fair or Poor condition classes, which was not particularly surprising because those streams were typically larger than the randomly selected sites. By their nature, larger streams are typically at lower elevations that often are more frequently exposed to historical or ongoing human disturbances. These disturbances were reflected by the calculated BMI metrics and MBII.

Given that this report represented the first year of data collection, there were few inferences or management recommendations that could be confidently made. Biological communities (including BMI) can vary through time due to a range of naturally occurring biotic phenomena (e.g., interspecific competition, predation) and abiotic disturbances (severe drought, floods). It will take several years to determine the degree to which BMI communities naturally vary

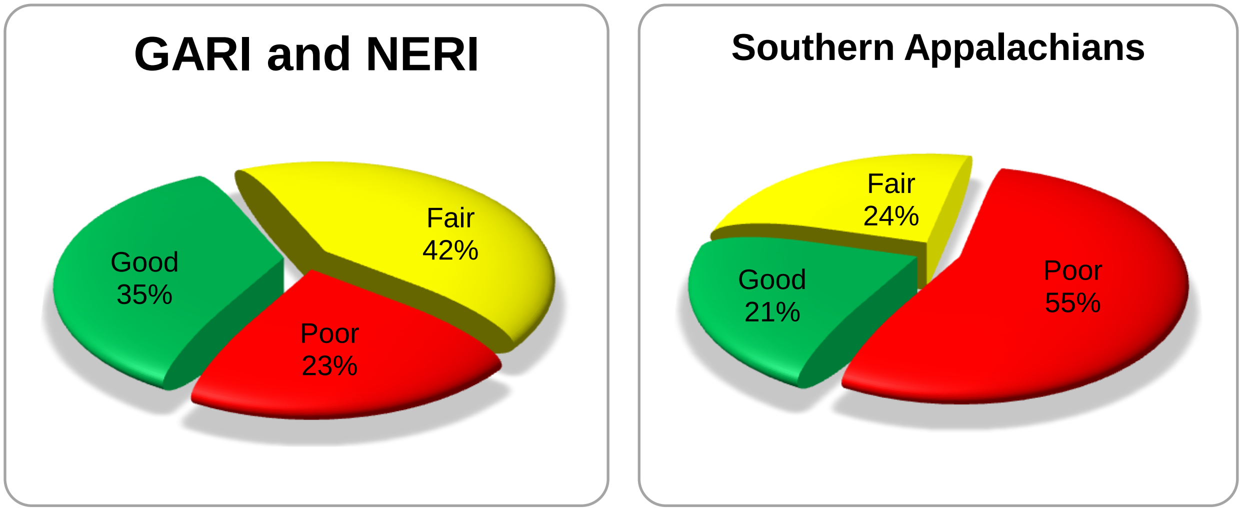

Figure 9. Condition class of randomly chosen wadeable streams throughout Gauley River National Recreation Area (GARI) and New River Gorge National River (NERI) in March 2009 (left) and Southern Appalachians Ecoregion (right, modified from USEPA 2006) based on Macroinvertebrate Biotic Integrity Index values. Percentages of GARI and NERI streams in each condition class are based on data from only probabilistically chosen sites.

throughout the ERMN. Once natural variability of BMI communities is quantified, we will be in a better position to make inferences about the relative condition of sampled streams.

With each future sampling season, the ERMN BMI monitoring program will be refined and improved. It is anticipated that metrics and indices will be calibrated so that more precise and accurate comparisons can be made among BLUE, GARI, and NERI streams and streams throughout the region. In addition to calibrating the MBII and its constituent metrics, we will add other measures of stream integrity as more data are gathered. For example, another meaningful way to express BMI community condition is with Observed/Expected Indices that estimate the number of taxa (e.g., genera) that have been lost (i.e., extirpated) from a given stream (Yuan 2008). To use these methods, the expected number of taxa for a given stream type must be established from the least-disturbed streams in the region. This process will likely begin after next season when assessments regarding natural variability of BMI communities can be at least coarsely made. During the next several years, we plan to cooperate with researchers from the Pennsylvania State University to standardize ERMN data to stream condition thresholds established during the WSA. That effort will allow more confident comparisons to be made between ERMN streams and similar streams throughout the ecoregion.

Literature Cited

Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, Second Edition. EPA 841-B-99-002. U.S. Environmental Protection Agency. Office of Water. Washington, DC.

Bode, R. W., M. A. Novak, L. E. Abele, D. L. Heitzman, and A. J. Smith. 2002. Quality assurance work plan for biological stream monitoring in New York state. Stream Biomonitoring Unit, Bureau of Water Assessment and Management, New York State Department of Environmental Conservation. Albany, NY. 41 pages.

Brower, J. E., and J. H. Zar. 1984. Field and laboratory methods for general ecology (2d ed.): Dubuque, IA. Wm. C. Brown Publishers. 226 pp.

Carter, J. L., and V. H. Resh. 2001. After site selection and before data analysis: sampling, sorting, and laboratory procedures used in stream benthic macroinvertebrate monitoring programs by USA state agencies. Journal of the North American Benthological Society 20(4):658–682.

Fancy, S. G., J. E. Gross, and S. L. Carter. 2009. Monitoring the condition of natural resources in US national parks. Environmental Monitoring and Assessment 151:161–174.

Herlihy, A. T., S. G. Paulsen, J. Van Sickle, J. L. Stoddard, C. P. Hawkins, and L. L. Yuan. 2008. Striving for consistency in a national assessment: the challenges of applying a reference-condition approach at a continental scale. Journal of the North American Benthological Society 27(4):860–877.

Klemm, D. J., K. A. Blocksom, F. A. Fulk, A. T. Herlihy, R. M. Hughes, P. R. Kaufmann, D. V. Peck, J. L. Stoddard, W. T. Thoeny, M. B. Griffith, and W. S. Davis. 2003. Development and evaluation of a macroinvertebrate biotic integrity index (MBII) for regionally assessing mid-Atlantic highlands streams. Environmental Management 31(5):656–669.

Marshall, M. R., and N. B. Piekielek. 2007. Eastern Rivers and Mountains Network Ecological Monitoring Plan. Natural Resource Report NPS/ERMN/NRR—2007/017. National Park Service. Fort Collins, CO.

Mattsson, B. J., and M. R. Marshall. 2009. Streamside Bird Monitoring Protocol for the Eastern Rivers and Mountains Network. Technical Report NPS/ERMN/NRR—2009/DRAFT. National Park Service. Fort Collins, CO.

Moulton, S. R., J. L. Carter, S. A. Grotheer, T. F. Cuffney, and T. M. Short. 2000. Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory – processing, taxonomy, and quality control of benthic macroinvertebrate samples. U.S. Geological Society Open-File Report 00-212. 61 pp.

Moulton, S. R., J. G. Kennen, R. M. Goldstein, and J. A. Hambrook. 2002. Revised protocols for sampling algal, invertebrate and fish communities as part of the National Water-Quality Assessment Program: U.S. Geological Survey Open-File Report 02-150. 87 pp.

Perles, S. J., J. C. Finley, and M. R. Marshall. 2009. Vegetation and soil monitoring protocol for the Eastern Rivers and Mountains Network. Natural Resource Report NPS/ERMN/NRR—2009/DRAFT. National Park Service. Fort Collins, CO.

Tzilkowski, C. J., A. S. Weber, and C. P. Ferreri. 2009. Benthic macroinvertebrate monitoring protocol for wadeable streams in the Eastern Rivers and Mountains Network. Natural Resource Report NPS/ERMN/NRR—2009/DRAFT. National Park Service. Fort Collins, CO.

United States Environmental Protection Agency (USEPA). 2006. Wadeable streams assessment: a collaborative survey of the nation's streams. EPA 841-B-06-002. Office of Research and Development. Office of Water. Washington, DC. 113 pp.

Yuan, L. L., C. P. Hawkins, and J. Van Sickle. 2008. Effects of regionalization decisions on an O/E index for the national assessment. Journal of the North American Benthological Society 27:892–905.

The Department of the Interior protects and manages the nation's natural resources and cultural heritage; provides scientific and other information about those resources; and honors its special responsibilities to American Indians, Alaska Natives, and affiliated Island Communities.

NPS 619/101086, 600/101086, 637/101086, February 2010

National Park Service
U.S. Department of the Interior

Natural Resource Program Center
1201 Oakridge Drive, Suite 150
Fort Collins, CO 80525

www.nature.nps.gov